Abstract of the Disclosure

A method for objective speech quality assessment that accounts for phonetic contents, speaking styles or individual speaker differences by distorting speech signals under speech quality assessment. By using a distorted version of a speech signal,

5   it is possible to compensate for different phonetic contents, different individual speakers and different speaking styles when assessing speech quality. The amount of degradation in the objective speech quality assessment by distorting the speech signal is maintained similarly for different speech signals, especially when the amount of distortion of the distorted version of speech signal is severe. Objective speech quality assessment for the

10   distorted speech signal and the original undistorted speech signal are compared to obtain a speech quality assessment compensated for utterance dependent articulation.
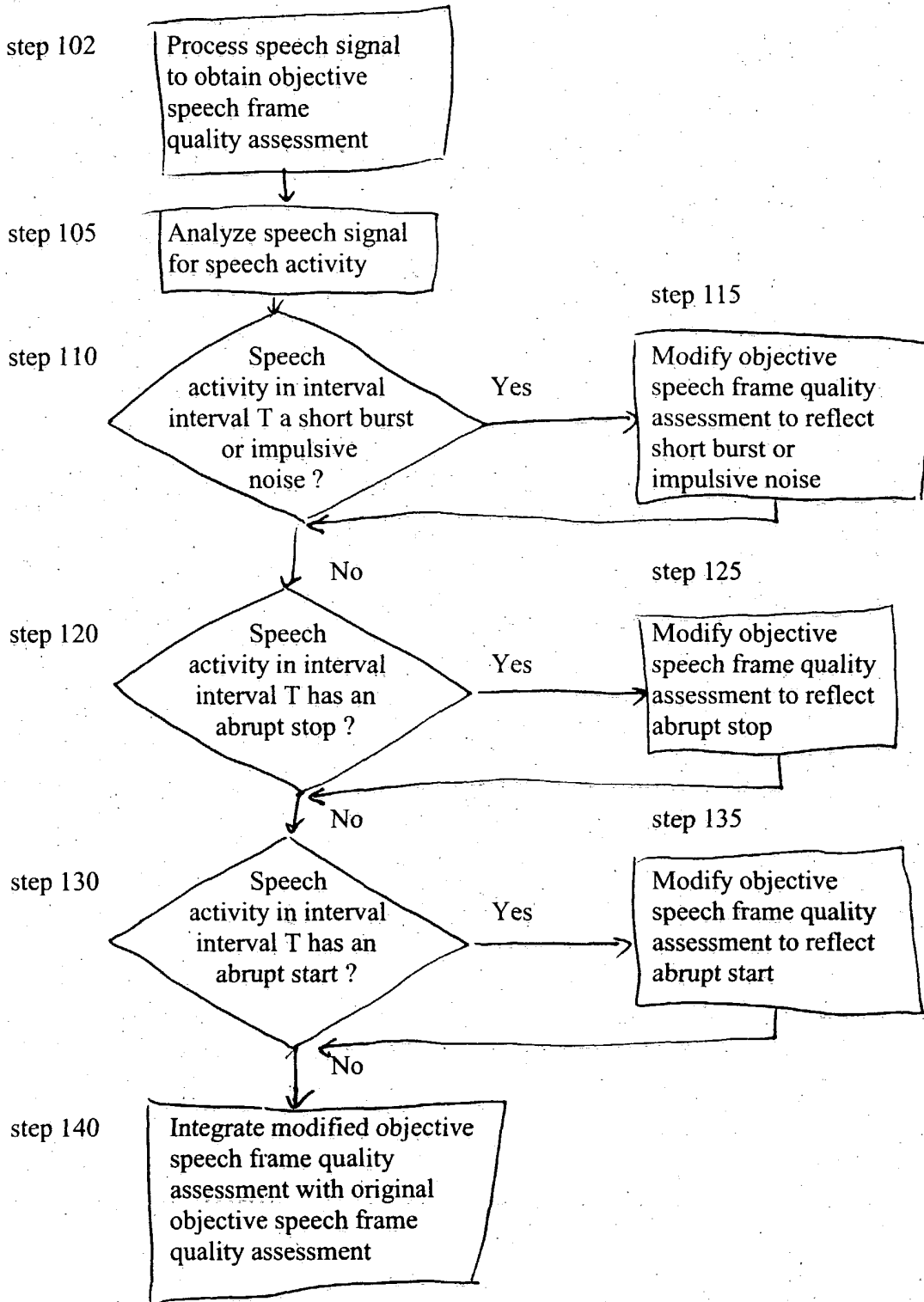
100

step 102 | Process speech signal to obtain objective speech frame quality assessment

step 105 | Analyze speech signal for speech activity

step 110 | Speech activity in interval interval T a short burst or impulsive noise ?

Yes →

step 115 | Modify objective speech frame quality assessment to reflect short burst or impulsive noise

No

step 120 | Speech activity in interval interval T has an abrupt stop ?

Yes →

step 125 | Modify objective speech frame quality assessment to reflect abrupt stop

No

step 130 | Speech activity in interval interval T has an abrupt start ?

Yes →

step 135 | Modify objective speech frame quality assessment to reflect abrupt start

No

step 140 | Integrate modified objective speech frame quality assessment with original objective speech frame quality assessment

Figure 1

200

step 205 — Sum envelope signals for all cochlear channels

step 210 — Determine frame envelopes using summed envelope signal and Hamming window

step 215 — Perform flooring operation on frame envelopes

step 220 — Determine time derivatives for floored frame envelopes

step 225 — Determine presence of voice activity using determined time derivatives
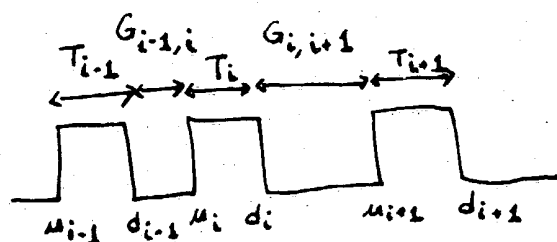
step 230 — Refine voice activity determination

Figure 2

$G_{i-1,i}$ $G_{i,i+1}$

T_{i-1} T_i T_{i+1}

u_{i-1} d_{i-1} u_i d_i u_{i+1} d_{i+1}

Figure 3

400

step 405 — Determine impulsive noise frame l_I for speech activity interval T

step 410 — Frame energy for impulsive noise frame l_I > noise threshold ?

Yes → step 415 — Speech activity interval T not short burst or impulsive noise

No ↓

step 420 — Interval T ≥ perception threshold and ≤ short burst threshold ?

Yes → step 425 — Speech activity interval T not short burst or impulsive noise

No ↓

step 430 — Maximum delta frame envelope > abrupt change threshold ?

Yes → step 435 — Speech activity interval T not short burst or impulsive noise

No ↓

step 440 — Frame m_I sufficiently annoying to human listener ?

Yes → step 445 — Speech activity interval T not short burst or impulsive noise

No ↓

step 445 — Human speech ?

Yes → step 455 — Speech activity interval T not short burst or impulsive noise

No ↓

step 460 — Modify objective speech frame quality assessment

Figure 4

500

step 505   Determine abrupt stop frame l_M for speech activity interval T

step 510   Delta frame energy for abrupt stop frame l_M > abrupt stop threshold ?

Yes

step 515

Speech activity interval T does not have an abrupt stop

No

step 520   Duration of interval T long enough ?

Yes

step 525

Speech activity interval T does not have an abrupt stop

No

step 530   Maximum delta frame envelope > stop-energy threshold ?

Yes

step 535

Speech activity interval T does not have an abrupt stop

No

step 540   Modify objective speech frame quality assessment

Figure 5

600

step 605 — Determine abrupt start frame l_S for speech activity interval T

step 615

step 610 — Delta frame energy for abrupt start frame l_M > abrupt start threshold ?

→ Yes → Speech activity interval T does not have an abrupt start

↓ No

step 625

step 620 — Duration of interval T long enough ?

→ Yes → Speech activity interval T does not have an abrupt start

↓ No

step 635

step 630 — Maximum delta frame envelope > start-energy threshold ?

→ Yes → Speech activity interval T does not have an abrupt start

↓ No

step 640 — Modify objective speech frame quality assessment

Figure 6